UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/532,919	01/16/2006	Albert Kooiman	DE 020239	4947

24737 7590 04/12/2010

PHILIPS INTELLECTUAL PROPERTY & STANDARDS
P.O. BOX 3001
BRIARCLIFF MANOR, NY 10510

EXAMINER	

ART UNIT	PAPER NUMBER

DATE MAILED: 04/12/2010

Please find below and/or attached an Office communication concerning this application or proceeding.

Commissioner for Patents
United States Patent and Trademark Office
P.O. Box 1450
Alexandria, VA 22313-1450
www.uspto.gov

BEFORE THE BOARD OF PATENT APPEALS AND INTERFERENCES

Application Number: 10/532,919
Filing Date: January 16, 2006
Appellant(s): KOOIMAN, ALBERT

Thomas Kocovsky
Robert Sieg
Fay Sharpe LLP
The Halle Building, 5th Floor
Cleveland, OH 44115-1843
For Appellant

EXAMINER'S ANSWER

This is in response to the appeal brief filed 12/24/2008 appealing from the Office action mailed 07/01/2008.

The appeal brief is filed in the new format under the revised BPAI final rule before the effective date of the BPAI final rule. The Office published the BPAI final rule to amend the rules governing practice before the BPAI in ex parte patent appeals. See Rules of Practice Before the Board of Patent Appeals and Interferences in Ex Parte Appeals; Final Rule, 73 FR 32938 (June 10, 2008), 1332 Off. Gaz. Pat. Office 47 (July 1, 2008). However, the effective date for the BPAI final rule has been delayed. See Rules of Practice Before the Board of Patent Appeals and Interferences in Ex Parte Appeals; Delay of Effective and Applicability Dates, 73 FR 74972 (December 10, 2008). In the notice published on November 20, 2008, the Office indicated that the Office will not hold an appeal brief as non-compliant solely for following the new format even though it is filed before the effective date. See Clarification of the Effective Date Provision in the Final Rule for Ex Parte Appeals, 73 FR 70282 (November 20, 2008). Since the appeal brief is otherwise acceptable, the Office has accepted the appeal brief filed by appellant.

(1) Real Party in Interest

A statement identifying by name the real party in interest is contained in the brief.

(2) Status of Amendments After Final

The appellant's statement of the status of amendments after final rejection contained in the brief is correct.

(3) Grounds of Rejection to be Reviewed on Appeal

The appellant's statement of the grounds of rejection to be reviewed on appeal is correct.

(4) Evidence Relied Upon

5,765,130	Nguyen	6-1998
5,978,763	Bridges	11-1999
6,173,266	Marx et al.	1-2001
6,336,091	Polikaitis et al.	1-2002
6,505,155	Vanbuskirk	1-2003
6,754,310	Steinbrenner et al.	6-2004
7,069,221	Crane et. al.	6-2006

Gerven, Stephen Van. "A Comparative Study of Speech Detection Method" Proc. 5th Eur. Conf. Speech Communications Technology Eurospeech '97, Rhodes, Greece 1997.

(5) Grounds of Rejection

The following ground(s) of rejection are applicable to the appealed claims:

Claim Rejections - 35 USC § 102

The following is a quotation from 35 U.S.C. 102:

A person shall be entitled to a patent unless –

> (b) the invention was patented or described in a printed publication in this or a foreign country or in public use or on sale in this country, more than one year prior to the date of application for patent in the United States.

Claims 1, 2, 4, and 10 are rejected under 35 U.S.C. 102(b) as anticipated by Audrius Polikaitis, et al, U.S. Patent Number 6336091 (hereinafter "Polikaitis, et al).

a. As to claims 1 and 10, Polikaitis teaches a method for operating a speech recognition system (Polikaitis, et al, Fig. 2 and Fig. 3),

detecting a speech signal of a user (Polikaitis, et al, Fig. 2 and Fig. 3, 215);

analyzing the speech signal to recognize speech information (Polikaitis, et al, Fig. 2 and Fig. 3, 220) contained in the speech signal,

determining a reception quality value or a noise value (Polikaitis, et al, Fig. 2 and Fig. 3, variables in 230, 240, 250, 260, the parameters used by the reference use energy calculations to determine reception quality of the voice from the user) which represents a current reception quality, and

switching the speech recognition system over to a mode of operation which is less sensitive to noise when the noise value exceeds a noise threshold or outputting an alert signal to the user (Polikaitis, et al, Fig. 2 and Fig. 3, 233, 243, 253, 263) when the reception quality value drops below a given reception quality threshold (Polikaitis, et al, Fig. 2 and Fig. 3, thresholds in 230, 240, 250, 260) or both.

b. In the context of this claim, the "or" being satisfied if the speech recognition system (1) "switches over to a mode of operation which is less

sensitive to noise" or (2) "outputs an alert signal to the user". Polikaitis, et al, teaches the invention of claim 1 consistent with option (2).

As to claim 2, Polikaitis teaches further comprising:

automatically resetting the speech recognition system to a previous mode of operation when the reception quality value (SQ) exceeds the reception quality threshold or when the noise value drops below the noise threshold (see Figure 3, output of element 230, input to repeat prompt to user 270 when it exceeds a second threshold user is prompted to repeat voice instruction during analysis window (see col. 6, lines 62-65).

As to claim 4, Polikaitis teaches a wherein

the reception quality value (S0) or the noise value is determined with a voice activity detector (see col. 4, lines 32-35, speech noise classifier (interpreted to be the speech/noise classifier in conjunction with microprocessor 110) used for determining noise and speech frames and col. 4, lines 42-col. 5, lines 15, features extracted as a result of the classification and see Figure 3, where the values are compared to a threshold to determine if speech recognition may be correct or incorrect.).

Claim Rejections - 35 USC § 103

Claim 3 is rejected under 35 U.S.C. 103(a) as being unpatentable over Polikaitis, et al, as applied to claim 1, and further in view of Nguyen, John N., USPN 5765130

(hereinafter "Nguyen") and Crane, Matthew, et al, USPN 7069221 (hereinafter "Crane, et al").

a) Polikaitis, et al, teaches the method of claim 1 upon which claim 3 depends. Polikaitis, et al, don't teach that a barge-in mode of operation is disabled based on the confidence measures hypothesis test results.

b) Nguyen teaches deactivating a barge-in mode operation of the speech recognition when the reception quality value drops below the reception quality threshold or the noise value exceeds the noise threshold (Nguyen, Col. 5 27-32), the barge-in mode of operation (Nguyen, Abstract) of the speech recognition system (Nguyen, Abstract).

c) It would have been obvious to one of ordinary skill in the art to implement the teachings of Nguyen, et al, into the teachings of Polikaitis, et al, since Polikaitis suggests a speech recognition system that performs confidence measures hypothesis testing via thresholding on the received signal to provide voice prompts to the user to manage the speech recognition process to be more effective and since Nguyen teaches a barge-in feature that prevents users' voices and barge-in echo from combining to disrupt the interactive session, also threshold-based, wherein his invention "provide[s] a method and apparatus for implementing ...barge in" in the analogous art of telephone voice recognition systems and where "...the invention is useful even in the absence of local echo cancellation, since it still provides a dynamic threshold for determination of

whether a user signal is being input concurrent with a prompt." (Nguyen, Col 7. lines 6-9).

d) Crane, et al, teaches a barge-in feature for a speech recognition system in which, when the source of signal is determined to be a non-target barge-in, the barge-in mode of operation is deactivated (Crane, et al, Fig. 3 Elements 70, 72, 74).

e) It would have been obvious to someone of ordinary skill in the art at the time the invention was made to implement the teachings of Crane, et al, into Polikaitis, et al, since Polikaitis, et al, teaches an automatic interactive speech recognition system that uses voice prompts with the user in combination with incoming signal conditions hypothesis testing to threshold for disruptions, and since Crane, et al, teaches a method of determining of whether a potential barge-in signal energy detected is that of a user, with thresholding of a confidence measure to enable or disable prompt play in a barge-in system as "in one embodiment, [wherein] recognizer 37 determines whether the sound received is a target or a non-target signal by obtaining a score for that signal, and determining whether the score exceeds a threshold for recognizing the signal as a target (or as a non-target) signal." (Crane, et al, "Description of Preferred Embodiments" ¶ 18).

Claim 5 is rejected under 35 U.S.C. 103(a) in light of Polikaitis, et al, as applied to claim 1 and further in view of S Van Gerven, and F Xie - Proc. Eurospeech, 1997 (hereinafter "Gerven and Xie").

f) Polikaitis, et al, teaches the method of claim 1 upon which claim 3 depends. Polikaitis, et al, don't teach that the estimates of noise level are based on measuring the background signal, i.e. the input signal when before the user speaks or during speech pauses.

a) Gerven and Xie teach that correct voice activity detection should include characterizing the noise during noise periods and characterizing the speech during speech periods. Gerven and Xie teach characterizing a reception quality value (Gerven and Xie, §A ¶ 1, "energy of the total signal in the presence of speech") or a noise value (Gerven and Xie, §A ¶ 1, "varying noise level"), determined on the basis of a background signal (Gerven and Xie, §A ¶ 1, "background noise") which is received prior to the beginning of the utterance and/or in a speech pause of the user (Gerven and Xie, §B ¶ 1, "voice inactive segments.") or both.

b) It would have been obvious for one of ordinary skill in the art the time to implement the teachings of Gerven and Xie into the teachings of Polikaitis, et al, since Polikaitis, et al, suggests the benefits of measuring signal metrics associated with speech and noise against thresholds to reduce operational error and Gerven and Xie teach obtaining signal and noise metrics during speech and non-speech periods, respectively. (Gerven and Xie, page 1, ¶ 4) ("Adaptive speech enhancement algorithms typically behave completely different during speech periods than during noise periods. During speech periods the algorithms should learn as much as possible about the speech source and during noise

periods as much as possible about the noise source(s). Correct voice activity detection (VAD) is therefore crucial to their success."). Gerven and Xie also suggest that measuring the background noise during speech pauses are the "classic energy threshold method" of voice activity detection. (Gerven and Xie, §**B** ¶ 1).

Claim 6 is rejected under 35 U.S.C. 103(a) as being obvious over Polikaitis, et al, as applied to claim 4 above further in light of Marx, Matthew, et al, USPN 6173266 (hereinafter "Marx, et al").

a) Polikaitis, et al, teaches the method of claim 1 upon which claim 3 depends. Polikaitis, et al, don't teach that a reception corruption signal is sent to a dialog control device.

b) Dialog control devices existed and were widely used in the art at the time the invention was made for the purposes of managing voiced interactive sessions in automatic speech recognition systems.

c) Marx, et al, teach a dialog control module feature for an automatic interactive speech recognition system (Marx, et al, Fig 4) characterized in that the voice activity detector (Marx, et al, Fig 2) applies the reception quality value (Marx, et al, Fig 2, 260) or the noise value (Marx, et al, Fig 2, 270) itself (Marx, et al, Fig 2, 250) and/or, when the reception quality value drops below the reception quality threshold (Marx, et al, Fig 2, 280) or when the noise value exceeds the noise

threshold (Marx, et al, Fig 2, 280), a reception corruption indication signal (Marx, et al, Fig 2, 215) to a dialog control device (Marx, et al, Fig 4, 430).

d) Thus, it would have been obvious for one of ordinary skill in the art at the time the invention was made to implement the teachings of Marx, et al, into the teachings of Polikaitis, et al, since Polikaitis, et al, suggest the benefits of an interactive speech recognition system with voice activity detection for confidence measures hypothesis testing and use the results to mitigate errors due to reception conditions, and since Marx, et al, suggest the use of dialog control modules to manage the user interaction session in conjunction with confidence measures hypothesis testing to manage the interactive session. (Marx, et al, Col. 3 Line 46).

Claim 7 is rejected under 35 U.S.C. 103(a) as being unpatentable over Polikaitis, et al, and Marx, et al, as applied to claim 1 above further in view of Vanbuskirk, et al, USPN 6505155 (hereinafter "Vanbuskirk, et al").

a) Polikaitis, et al, teach a speech recognition system with voiced error prompts to the user based on the results of confidence measures hypothesis testing of incoming signal features. Marx, et al, teaches a speech recognition system wherein a dialog control device manages the voiced interactive session with a user in an interactive, automated speech system. Polikaitis, et al, and Marx, et al, do not teach that the user is sent any particular information when hypothesis tests fail.

b) Vanbursick, et al, teaches a method for operating speech recognition system an incoming signal (Vanbuskirk, et al, Fig. 4A Element 22) is analyzed (Vanbuskirk, et al, Fig. 4A Element 25, 33) as regards the type of disturbance causing (Vanbuskirk, et al, Fig. 4A Element 29) the reception quality value to be below the reception quality threshold or the noise value to be above the noise threshold (Vanbuskirk, et al, Fig. 4A Element 31), and that the dialog control device initiates the output of a prompt to the user (Vanbuskirk, et al, Fig. 4C-G Element 49, 44, 51,53, 57) who is thus given the information that the reception conditions are poor (Vanbuskirk, et al, Fig. 4C-G Element C) ."

c) It would have been obvious to one of ordinary skill in the art at the time the invention was made to implement the teachings of Vanbursick, et al, into the speech recognition system taught in Polikaitis, et al since Polikaitis suggests that a voiced alert to the user that the voice recognition conditions are likely to lead to an error, while Vanbuskirk suggests his system of dynamically composed voice prompts to the user reflecting poor ambient noise conditions serves that purpose. (Vanbuskirk, et al, ¶¶ 15 and 21). (i.e. The invention serves to "anticipate that recognition errors in consequence of heightened background noise ... [and] ... proactively adjust feedback [to the user]." and "Responsive to predicted adequate recognition accuracy, the present invention could reduce prompt feedback in the computer responsive prompt.").

d) Vanbursick's invention provides for dynamically composed voiced prompts to the user based on confidence measure hypothesis testing of the received signals.

Vanbuskirk does not teach explicitly that the information of the dynamically composed voice prompt to the user is that the reception conditions are poor when the confidence measures test fails.

e) The sending of information about reception condition problems is already anticipated in Polikaitis: "Alternatively, the microprocessor may permit the speech recognition processing to continue with a warning that the speech recognition output may be incorrect due to the error in the speech signal format" wherein errors in the speech signal format include "speech energy, noise energy, start energy, end energy, the percentage of clipped speech samples and other speech or signal related parameters within the speech acquisition window." (Polikaitis, et al, Col 2. Lines 41 and 51).

Claim 8 is rejected under 35 U.S.C. 103(a) as being unpatentable over Polikaitis, et al, as in claim 1 above further in view of Vanbuskirk, et al. and Steinbrenner, Kurt W., et al, USPN 6754310 (hereinafter "Steinbrenner, et al").

a) Polikaitis, et al, teaches the method of claim 1 upon which claim 3 depends. Polikaitis, et al, does not teach that the incoming signal is analyzed for the type of reception problem that occurs when the confidence measures hypothesis testing fails, and that this information is provided to the user via voiced messages.

b) Steinbrenner, et al, teach a method for operating an interactive automatic telephony system wherein analyzing an incoming signal (Steinbrenner, et al, Col.

2 Lines 31-42) is analyzed (Steinbrenner, et al, Col. 6 line 64 – Col. 7 line 1) as regards the type of disturbance causing the reception quality value to be below the reception quality threshold or the noise value to be above the noise threshold (Steinbrenner, et al, Col. 6 line 67), and that a prompt (Steinbrenner, et al, Fig. 5. Element 80; Fig. 7, 124) which contains diagnostic information (Steinbrenner, et al, Fig. 5. Element 78; Fig. 7, 122) is and outputting a prompt (Steinbrenner, et al, Fig. 5. Element 82; Fig. 7, 126) to the user.

c) Also, further note that Applicant's preferred embodiments of his invention included telephone-based systems such as those in Steinbrenner, et al: "Examples of such speech dialog systems are automatic answering and information systems which nowadays are used in particular by some large companies and public services so as to offer a caller as quickly and as comfortably as possible with the desired information ... [f]urther examples in this respect are automatic telephone information systems..." (Specification, ¶ 2) In the embodiments of Applicant's invention that consist of interactive telephone answering and information systems, the diagnostic information to be provided back to the user would necessarily be that taught in Steinbrenner, et al.

d) It would have been obvious to a person of ordinary skill in the art at the time the invention was made to implement the teachings of Steinbrenner, et al, into Polikaitis, et al, since Polikaitis, et al, suggests voiced alert prompts to the user under poor reception conditions while Steinbrenner, et al, describes the benefits of combining voiced prompts to the user containing network and device

diagnostic information in the analogous art of interactive automatic telephony. (Steinbrenner, et al, Col. 3 Lines 25 – 29)

Claims 9 and 11-13 are rejected under 35 U.S.C. 103(a) as being unpatentable over Polikaitis, et al, further in view of Marx, et al, and Bridges, James USPN 5978763 (hereinafter "Bridges").

a) As to claim 9, Polikaitis teaches a speech recognition system (Polikaitis, et al, Fig. 2 and Fig. 3) for the detection of a speech signal of a user (Polikaitis, et al, Fig. 2 and Fig. 3, 215) and a speech recognition device (Polikaitis, et al, Fig. 2 and Fig. 3, 290) so as to recognize speech information contained in the speech signal, characterized in that it comprises a quality control device for determining a reception quality value or a noise value, (Polikaitis, et al, Fig. 2 and Fig. 3, variables in 230, 240, 250, 260) representing a current reception quality, a comparator for comparing the reception quality value with a predetermined reception quality threshold or for comparing the noise value with a given noise threshold (Polikaitis, et al, Fig. 2 and Fig. 3, thresholds in 230, 240, 250, 260), and control means which are constructed in such a manner that the speech recognition system is switched over to a mode of operation which is less sensitive to noise and/or an alert signal is output to the user (Polikaitis, et al, Fig. 2 and Fig. 3, 233, 243, 253, 263) when the reception quality value drops below a given reception quality threshold or when the noise value exceeds a noise threshold.

b) In the context of this claim, the "or" being satisfied if the speech recognition system (1) "switches over to a mode of operation which is less sensitive to noise" or (2) "outputs an alert signal to the user". Polikaitis, et al, teaches the invention of claim 1 consistent with option (2).

c) Polikaitis, et al, do not teach that a control means causes the speech recognition system to send an alert signal to the user.

d) Marx, et al, teach a dialog control module that provides a control means for controlling an automatic interactive speech recognition session (Marx, et al, Fig 4) characterized in that the voice activity detector (Marx, et al, Fig 2) applies the reception quality value (Marx, et al, Fig 2, 260) or the noise value (Marx, et al, Fig 2, 270) itself (Marx, et al, Fig 2, 250) and/or, when the reception quality value drops below the reception quality threshold (Marx, et al, Fig 2, 280) or when the noise value exceeds the noise threshold (Marx, et al, Fig 2, 280), a reception corruption indication signal (Marx, et al, Fig 2, 215) to a dialog control device (Marx, et al, Fig 4, 430).

e) It would have been obvious for one of ordinary skill in the art at the time the invention was made to implement the teachings of Marx, et al, into the teachings of Polikaitis, et al, since Polikaitis, et al, suggest the benefits of an interactive speech recognition system with voice activity detection for confidence measures hypothesis testing to mitigate errors due to reception conditions, and since Marx, et al, suggest the use of dialog control modules to manage the user interaction

session in conjunction with confidence measures hypothesis testing to manage the interactive session. (Marx, et al, Col. 3 Line 46).

f) Polikaitis, et al, does not teach that a comparator is used for thresholding. However, a device that tests a value against a threshold is a kind of comparator. Thus, Applicant adds no new limitation in claim language where an initially narrow limitation (testing against a threshold) is followed by a broader limitation (use of a comparator).

g) Embodiments in which the dependent claim's added limitation of the comparator does not imply a limitation that is broader than the independent claim's thresholding includes applications in which the thresholding is a function that is separate and apart from the comparison against a threshold.

h) Bridges teaches a voice activity detection method that uses a comparator (Bridges, Fig 2, 268) to compare a received signal against an adaptive threshold (Bridges, Fig. 2) in a voice activity detector (also Bridges, ¶ 18).

i) Thus, it would have been obvious to one of ordinary skill in the art at the time the invention was made to implement the teachings of Bridges into the teachings of Polikaitis, et al, since Polikaitis, et al, suggests a speech recognition system that tests received signal for quality measures against a threshold and Bridges suggests that the use of a "threshold comparator" improves the performance of the voice activity detection in the case where echo return loss interferes with voice prompt system performance. (Bridges, ¶ 20) ("Controlling the threshold on the basis of the echo return loss measured not only reduces the number of false

triggering by the voice activity detector due to echo, but also reduces the number of triggering of the voice activity detector when the user makes a response over a line having a high amount of echo.").

As to claim 11, Polikaitis in view of Marx *et al.* in view of Bridges teach all of the limitations as in claim 9.

Furthermore, Polikaitis teaches a voice activity detector (see col. 4, lines 32-35, speech noise classifier (interpreted to be the speech/noise classifier in conjunction with microprocessor 110) used for determining noise and speech frames and col. 4, lines 42-col. 5, lines 15, features extracted as a result of the classification and see Figure 3, where the values are compared to a threshold to determine if speech recognition may be correct or incorrect.)

As to claim 12, Polikaitis in view of Marx *et al.* in view of Bridges teach all of the limitations as in claim 9.

Furthermore, Marx, et al, teach a control means for controlling an automatic (Marx, et al, Fig 4, dialogue modules are presented to the telephony interface based on speech input and output) further comprises a barge-in switching unit (see col. 7, lines 20-28, software to detect caller speech to provide barge-in detection and handling).

As to claim 13, Polikaitis in view of Marx *et al.* in view of Bridges teach all of the limitations as in claim 9.

Furthermore, Marx teaches. wherein the control means further comprises a dialog control device (see Figure 4, all components, dialogue modules are used to interface with telephony systems to control speech input and output and see col. 6, lines 62-67).

(6) Response to Argument

Claims 1, 4, and 10 are Rejected under 35 U.S.C. §102 as being anticipated by Polikaitis

Appellant asserts on pages 14-18

> In rejecting claim 1, the Office Action directs Applicants to the various thresholds in 230, 240,250 and 260 for the reception quality threshold. While there are thresholds disclosed, none are reception quality thresholds or a noise threshold as claimed. Rather there are start energy thresholds and end energy thresholds, which relate to energy in frames. As such, the applied art fails to disclose at least one feature of claim 1. As a result, a prima facie case of anticipation has not been established. Therefore, claim 1 and the claims that depend therefrom are patentable over the applied art...
>
> Thus, the decisions or determinations made by Polikaitis all relate to user errors and do not relate to determining a value which represents a current reception quality in the sense of strength of the cell phone or other incoming signal, noisy environment of the user, or other factors other than user error.

In response to the Appellant's arguments that the reception quality thresholds or a noise threshold are not disclosed by the thresholds in Polikaitis, the Examiner cannot concur. Polikaitis, in Figure 3, teaches the use of thresholds, which utilizes energy

values, noise energy, and other parameters to indicate the quality of the input signal received from the user from a communication device. In other words, the reception quality of the speech input from the user at the microphone is being evaluated by the use of the thresholds. Thus, the thresholds of Polikaitis are reception quality thresholds as they are used to evaluate the received signal quality at the microphone (see Polikaitis col. 2, lines 15-27) from a user for use in speech recognition. Further, the Appellant asserts that the reception quality threshold is determined from telephone links, reception channel, or background noise level. However, these represent various examples of what the reception quality threshold is determined from. The Appellant in the specification has not provided a special definition for the “reception quality threshold". Further, Polikaitis does measure in one of the thresholds the background noise level in relation to user input. In Figure 3, step 260, col. 8, lines 47-51, the speech energy and the noise energy is compared to determine if more noise is present. If this occurs, the user is notified. Thus, the quality of received input with respect to background noise is evaluated by Polikaitis and outputted to the user. The current reception quality is evaluated by Polikaitis as the user enters new speech (see Polikaitis col. 9, lines 33-34, where the entire process in Figure 3 repeats for new speech which is input.

Appellant asserts on page 18

> Claim 1 further calls for switching the speech recognition system over to a mode of operation which is less sensitive to noise when the noise value exceeds a noise threshold or outputting an alert signal to the user when the reception quality value drops below a given reception quality threshold. When Polikaitis

> determines that one of Error 1-Error 4 has occurred, Polikaitis merely prompts the user and/or adjusts the window size or the level of amplification. Polikaitis does not switch over to a mode of operation which is less sensitive to noise. Error 1 and Error 2 are not related to noise. With Error 3, if the amplification is adjusted downward, this is not a different mode of operation and it is more sensitive to noise. In Error 4, in which the user is speaking too quietly, Polikaitis increases the amplification which also amplifies the noise. Adjusting the amplification is not changing to a different mode of operation, much less a mode which is less sensitive to noise. Higher amplification would make the Polikaitis system more sensitive to noise.

In response to the Appellant's arguments that claim 1 further calls for switching the speech recognition system over to a mode of operation less sensitive to noise, the Examiner cannot concur. As indicated in the Office Actions dated 12/12/2007 and 07/01/2008, the claims use "or" terminology. Thus, the claims are satisfied if the prior art either (1) "switched over to a mode of operation which is less sensitive to noise" or (2) "outputting an alert signal to the user...." The teachings of Polikaitis anticipate option (2). This is described in Figure 2 and 3, informing user 233, 243, 253, and 263 and col. 6, lines 49-52, col. 7, lines 36-39, col. 8, lines 16-19,and lines 59-61.

Appellant further asserts on page 18

> Polikaitis does not determine a reception quality value, as discussed above. Rather, Polikaitis determines user error. If the Polikaitis user commits one of Errors 1-4, Polikaitis provides instructions how to cure the error. Instructing a user how to cure user error does not meet the limitation of outputting an alert signal which alerts the user that the reception quality value is below a given quality threshold.

In response to the Appellant's arguments that Polikaitis does not determine a reception quality value, the Examiner cannot concur. As stated above, Polikaitis does

teach a reception quality threshold. The Appellant has provided a broad definition in their specification. Thus, although Polikaitis may be directed to correction of user error, the subsequent input into the system is being changed as a result of the quality of input initially received by the system. This quality is evaluated using the four thresholds (see Polikaitis Figures 2 and 3). Therefore, Polikaitis does teach the determination of a reception quality value which is determined through the result of each threshold comparison steps 230, 240, 250, and 260 with corresponding alerts 233, 243, 253, and 263.

Appellant asserts on page 19

> Claim 2 calls for automatically setting a speech recognition system to a previous mode of operation. Polikaitis does not change modes of operation. At best, Polikaitis lengthens the speech acquisition window, increases the amplification or decreases the amplification, but does not change to a different mode. Adjusting the length of a speech acquisition window or the amount of amplification is not resetting to a previous mode of operation.
>
> Moreover, claim 2 calls for the resetting to the previous mode to be when the reception quality value exceeds a reception quality threshold or the noise value drops below" the noise threshold. Polikaitis extends the speech acquisition window in response to user error. Even if adjusting the length of the acquisition window were to be considered a mode, which it is not, Polikaitis does not describe changing the length of the user window in response to a reception quality value exceeding a quality threshold. First, speaking over the start or the end of the window is not a reception quality value. Second, speaking or not speaking over the beginning or end of the window is not a condition that exceeds or does not exceed a threshold. Analogously, there is no suggestion in Polikaitis of resetting mode of operation in response to noise dropping below a noise threshold.
>
> Polikaitis does not set forth criteria to returning the window length or amplification to their original length or amplitude. Returning to these original settings is not a change of mode and is not responsive to a threshold level being crossed.

In response to the Appellant's argument that Polikaitis does not change the modes of operation, the Examiner cannot concur. In Figure 3, and col. 6, lines 62-65, Polikaitis teaches the changing of modes in the cited sections. The claims are broadly written and do not specifically indicate how the speech recognition system is reset. For example, reset to a prior step with modification or reset with no modification (see Appellant's arguments directly above, 3rd paragraph). The Appellant's published specification, specifically paragraph [0012], does not describe how the system is reset. However, an example is given and is further stated that the user is a key factor in allowing the reception quality to become better This is similar to Polikaitis usage of the four thresholds and if not met re-prompting for user input to enhance reception quality of the input signal for speech recognition and to therefore satisfy the thresholds. The re-prompting goes to a previous mode of operation, which is the capture of user input via a microphone as seen and described in Figure 3, and col. 6, lines 62-65.

Further, the Appellant argues in the 2nd paragraph above that Polikaitis does not change the length of the user window in response to a reception quality threshold, the Examiner cannot concur. In Polikaitis, the return of the operation of requesting speech input from the user when value in Figure 3, steps 230 or 240 exceed a threshold (conditions for whether speech is spoken over the start of speech or end of speech window), then user is prompted to re-enter speech. The system is reset to a previous mode, specifically re-entry of speech from user. Further, speaking over the start or the end of the window is a reception quality value. Polikaitis as described in Figure 3, steps

230 and 240 as well as col. 6, lines 29-44 and col. 7, lines15-32 describes the usage of energy information in the received speech and comparing such values to a threshold. The result is a quality determination of received speech since if user speaks over the start of an acquisition window or over the end of the acquisition window, the received signal by the system is an incomplete waveform which will not be properly recognized (see Polikaitis coll. 1, lines 39-56).

Claim 3 is Rejected under 35 U.S.C. §103 under Polikaitis in view of Nguyen in view of Crane

Appellant asserts on pages 20 and 21

> Claim 3 calls for deactivating a barge-in mode of operation when the reception quality value drops below the reception quality threshold or the noise value exceeds the noise value threshold. First, Polikaitis has no barge-in mode to deactivate. Rather, Polikaitis is very specific to systems having a speech acquisition window as illustrated in Figure 4.
>
> Second, column 5, lines 27-34 of Nguyen referenced by the Examiner, contrary to the Examiner's assertion, does not deactivate a barge-in mode. Rather, column 5, lines 27-34 provides a description for how the Nguyen system tells that the user is barging-in. Nguyen is always in a barge-in mode. Thus, neither Polikaitis, which has no barge- in system, nor Nguyen which is always in a barge-in mode switch into or out of a barge-in mode nor do they teach or fairly suggest deactivating a barge-in system, much less what criteria should be used for such a deactivation.
>
> Nguyen would first need to inspire Polikaitis to add a barge-in system. Even if Nguyen would have so inspired Polikaitis, which it is submitted that he would not have, neither Polikaitis nor Nguyen inform nor put the reader in possession of the idea that one should from time to time deactivate a barge-in aspect, much less what criteria one would or should use to deactivate a barge-in aspect.

In response to the Appellant's argument in paragraph one that Polikaitis is very specific to having a speech acquisition window, the Examiner cannot concur. The system and method of Polikaitis is not limited to speech with speech acquisition window as asserted. In col. 6, lines 25-28, Polikaitis describes situations where the user speaks too loudly or softly, which is not based on a speech acquisition window but rather the amplitude. Furthermore, Polikaitis suggests a speech recognition system that performs various determinations of the received signal and the providing of voice prompts (see Figure 2 or 3, 270 based on the comparison of the parameters to a thresholds for speech recognition. Therefore, the inclusion of Nguyen of a barge-in mode within Polikaitis would have been obvious to one skilled in the art.

In response to the 2nd and 3rd paragraph above, the Examiner cannot concur. Nguyen does teach a barge-in detection that is based on a reception quality threshold. In col. 4, lines 46-67 and col.5, lines 9-21, the detection criteria for speech are described in relation to noise. Thus, speech is detected from the signal to determine if barge-in has occurred. This is done by evaluating with respect to a reception quality threshold of the input signal, Q1. Nguyen's system does inspire Polikaitis' teachings as it provides detection of barge-in and barge-in echo that may result in an interactive session (i.e. speech recognition). The system of Polikaitis would benefit from such teachings since it would enable the capture of speech information during the playing of prompt 270 to prevent the speech input being spoken over the start of the speech acquisition window and to allow for correct input into the speech recognizer upon subsequent input from the user. Furthermore, deactivation of barge-in was relied upon

Crane. Crane in Figure 3, step 70 determines whether a target source has been detected. If the condition is true barge-in is activated (i.e. step 72) and speech recognition takes place, if not, then barge-in is not activated. This specifically described in step 74, where the speech prompts continues or is re-played. Neither the claims nor in paragraph [0034] of the Appellant's published Specification does the deactivation have a special meaning. Rather, the deactivation of the barge-in enables the prompt not to be interrupted. Likewise, Crane, when a target signal is not identified, the prompt continues to be played and is not interrupted by a non-target signal. Claim 3 does not specify how the barge-in mode is deactivated. Thus, the claim can be interpreted as described above from the deactivation of barge-in for non-target signals as taught by Crane.

Claim 5 is Rejected under 35 U.S.C. §103 under Polikaitis in view of Gerven

Appellant asserts on pages 21 and 22

> Claim 5 calls for determining the reception quality value or the noise on the basis of a background signal which is received prior to a beginning of an utterance or in a speech pause of the user, or both. Gerven does not supply this shortcoming of Polikaitis. Rather, Gerven is concerned with determining when speech is present and when only background noise is present (Gerven, page 1, second paragraph). Gerven presents three different algorithms, Algorithm 1, Algorithm 2, Algorithm 3 for differentiating between noise and speech (Gerven, sections 2.1, 2.2, 2.3). Gerven compares the performance of these three algorithms (Gerven, section 3). While Gerven addresses, in detail, how to tell when speech is present, nowhere in Gerven is there any suggestion of analyzing the signal when no speech is present to determine either a reception quality value or a noise value for switching modes of operation. If Polikaitis were to consider the Gerven reference, Polikaitis might be motivated to select one of Gerven's Algorithms 1, 2, or 3, but Polikaitis would not be put in possession or be motivated to alter the Polikaitis system to analyse the background signal. Gervan dies not teach or fairly suggest the subject matter of claim 5.

In response to the Appellant's argument that nowhere in Gerven is there any suggestion of analyzing the signal when no speech is present to determine either a reception quality value or a noise value for switching modes of operation, the Examiner cannot concur. In Gerven, page 3, sect. 2.3, 2nd paragraph Gerven analyzes a signal and recalculates parameters during non-speech periods (i.e. when no speech is present). Further, Appellants assert that Polikaitis would not be motivated to analyze to a background signal. The Examiner cannot concur. In Polikaitis col. 4, lines 32-41, a speech/noise classifier is described for determining periods of speech and noise. Polikaitis uses the classification result to determine the speech energy and noise energy (see col. 5, lines 6-23) when a user speaks too softly (see col. 8, lines 46-51, where the ratio is a reception quality value used to determine if a user spoke softly). Similarly, Gerven teaches methods to detect speech when noise is present which would aide in the classification process of Polikaitis (see Gerven, page 1, sect. 1, 1st paragraph). Hence, one would be motivated to alter the Polikaitis system a specific speech detection algorithm as Polikaitis already analyses a signal for noise (background signal) and speech.

Claims 6 is Rejected under 35 U.S.C. §103 under Polikaitis in view of Marx and claim 7 is rejected under 35 U.S.C. §103 under Polikaitis in view of Marx in view of Vanbuskirk

Appellant asserts on pages 22 and 23

> Claim 6 calls for a reception corruption indication signal. Marx does not cure this shortcoming of Polikaitis. Steps 260, 270 of Marx determine a

> confidence level and if the confidence is low, a step 15 prompts the user with a prompt such as "I'm sorry, I didn't hear your response. Please repeat your answer now." (Marx column 2, lines 5-9 and 26-39). A lack of confidence flows from factors such factors as speaking too loud or too soft, accents, word choice, and the like. Thus, rather than a reception corruption indication signal, Marx merely determines a confidence level with which the response was interpreted. Thus, Claim 6 goes to reception; whereas, Marx goes to interpretation confidence. Accordingly, it is submitted that claim 6 and claim 7 dependent therefrom distinguish patentably over the references of record.

In response to the Appellant's argument that Claim 6 goes to reception; whereas, Marx goes to interpretation confidence, the Examiner cannot concur. As admitted by the Appellant, Marx uses confidence metrics to determine the user's quality of speech input. A lack of confidence results in a prompt to the user to repeat input. The determined confidence parameter generates a reception indication signal as to the quality of speech input. This is supported by Figure 2, step 280 and the output of step 280 to 290 (take appropriate action) and step 215 (re-prompt). Considering the latter, for example, the system (e.g. processor consisting of program modules and see Figure 4, dialogue modules 430) sends a control signal for playing the prompt via a speaker (see Figure 4, speech output components 440 through telephony interface 460, which inherently consists of a speaker) to indicate to the user the quality of speech input, where a re-prompt indicates low quality and an appropriate action indicates higher quality as such is only executed once the confidence is high (See col. 2, lines 36-39). The Appellant's specification is consistent with this interpretation. The Appellant's published specification paragraph [0018] describes the outputting of an indication signal to the dialog control device or other components when reception quality drops. There is no description as to what this indication signal represents and thus can be interpreted as a

signal determined as a result from confidence scoring for outputting a prompt or and action as taught by Marx.

Claims 8 is Rejected under 35 U.S.C. §103 under Polikaitis in view Vanbuskirk in view of Steinbrenner

Appellant asserts on pages 23 and 24

> Claim 8 calls for analyzing an incoming signal for a type of disturbance causing the reception quality value to be below the reception quality threshold or the noise value to be above the noise threshold. Steinbrenner is directed to a telephony interface device for providing status and diagnostic information for a telephone operatively coupled to a telephone interface device (Steinbrenner, column 1, lines 7-11). That is, Steinbrenner is concerned with telephone system diagnostic information, particularly when a telephone is off the hook, which provides problems to the network switching system. Adding the Steinbrenner type system to Polikaitis would merely tell Polikaitis when a phone is off the hook, or the like. However, if a telephone is off the hook, the Polikaitis system would never come into play. That is, if the user's telephone is off the hook and not in use, the user will not be calling in to the Polikaitis system so Polikaitis has no need for or use for the information that a telephone somewhere in the network is off the hook or otherwise experiencing network problems.

In response to the Appellant's argument that adding the Steinbrenner type system to Polikaitis would merely tell Polikaitis when a phone is off the hook and if a telephone is off the hook, and the Polikaitis system would never come into play, the Examiner cannot concur. In Steinbrenner col. 8, lines 17-24, specifically describes that the off-hook feature that relates to the user picking up a phone (e.g. off hook), which is connected to a telephony interface. Thus, the phone in Steinbrenner is in use and would enable calling within the Polikaitis system and would enable the conveying of actual information with respect to error (low quality) input from the user in order to prevent

future occurrences. Steinbrenner teaches this type of diagnostic information in col. 3, lines 35-42. Such teachings would benefit the system of Polikaitis as Polikaitis teaches the use of telephony devices (see Polikaitis col. 3, lines 6-8) as a communication device through which the user inputs speech and also teaches the playback of prompts. The incorporation of utilizing diagnostic information into the system of Polikaitis enables voiced prompts to the user to be conveyed about possible problems in the initial received speech.

Claims 9 and 11-13 are Rejected under 35 U.S.C. §103 under Polikaitis in view Marx in view of Bridges

In pages 25 and 26 of the Appeal Brief, the Appellant provides similar arguments as discussed with respect to claim 1 and are referred to above.

Appellant asserts on pages 26 and 27

> To the contrary, Marx does not disclose outputting an alert signal either that the reception quality has dropped below a threshold or that a noise value exceeds a threshold. Rather, elements 260, 270, 280, and 215 of Marx relate to whether an answer was understood. Marx suggests that if the system does not have sufficient confidence in its understanding of the answer, then step 215 should prompt the user such as "I'm sorry, I didn't hear your response. Please repeat your answer now." Thus, like Polikaitis, Marx is not making a determination whether reception quality is below" a threshold or whether noise exceeds a threshold nor outputs an alert signal indicative of low reception quality or high noise. Accordingly, it is submitted that Marx does not cure this shortcoming of Polikaitis.

In response to the Appellant's argument that Marx does not disclose outputting an alert signal either that the reception quality has dropped below a threshold or that a

noise value exceeds a threshold, the Examiner cannot concur. Polikaitis was used as the primary reference for teaching the outputting of the alert signal when the reception quality dropped below a threshold. This is described in Figures 2 or 3, elements 233, 243, 253, and 263 and col. 6, lines 49-52, col. 7, lines 36-39, col. 8, lines 16-19,and lines 59-61. The alert signal can comprise a vibration, visual output, or voice output. The feature that was relied upon in Marx was the control means. Figure 4 of Marx teaches such a feature where the dialog modules interface with the various speech output components and subsequently interface with the telephony interface. Further, in Figure 2, voice prompts are output at 215 when low confidence with respect to the received speech is obtained. This output of a re-prompt is analogous to the primary reference which alerts the user via voice prompts when a specific threshold is not met (see Polikaitis, Figure 2 or 3). The usage of the control means of Marx enables the alerts of Polikaitis to be sent to the user indicating reception quality issues with coordination of the prompts as well as the speech recognition system.

Appellant asserts on pages 27 and 28

> Bridges, as the Examiner notes in paragraph 17h of the Final Rejection, does disclose a comparator. However, the comparator 268 of Bridges is for a different purpose and produces a different result. Specifically, the comparator 268 of Bridges determines whether or not an incoming signal is direct speech to deactivate the speech generator and activate the speech recognizer (Bridges, column 6, lines 5-13). It is submitted that if one were to add the comparator 268 of Bridges to Polikaitis, that Bridge's comparator would be used in a part of the system which would arbitrate between whether the unshown interface is in a mode for the user to provide a speech input 215 or whether the interface is a mode in which the user is prompted 270 or informed 275. The bridges comparator would neither replace steps 230, 240, 250, or 260 of Polikaitis, nor

would replacing these steps with a comparator cure the shortcomings of Polikaitis noted above.

In response to the Appellant's argument that the comparator 268 of Bridges is for a different purpose and produces a different result, the Examiner cannot concur. The comparator that is taught in Bridges in Figure 2, 268 and col. 6, lines 5-13 teach a comparator that uses a threshold. The primary reference of Polikaitis teaches the use of reception quality thresholds in terms of speech input from a user. The component lacking in Polikaitis is a comparator component for performing the thresholding operations. The thresholding is performed by a microprocessor 110 (see Polikaitis col. 6, lines 29-31). Incorporating Bridge's comparator into the processing component for performing the various thresholding operations allow a specific component within the processing system to be realized to be realized (i.e. change or addition of software code for performing the thresholding (i.e. function call for performing the comparison)).It should be made evident that the functionality of bridge's comparator is not being incorporated in the system of Polikaitis but rather the comparator with its thresholding capabilities, where one skilled in the art would have sufficient knowledge to modify software to include the thresholding operations of Polikaitis in a comparator (i.e. function call).

Claims 12 is Rejected under 35 U.S.C. §103 under Polikaitis in view Marx in view of Bridges

Appellant asserts on pages 28 and 29

> Claim 12 calls for the control means to comprise a barge-in switching unit. First, it is submitted that requiring the barge-in switching unit now requires claim 11 to be interpreted in accordance with the Examiner's option (1) set forth in

> paragraph 17b of the Final Rejection. The Examiner does not assert and effectively concedes that Polikaitis does not disclose or teach switching modes.
>
> On page 18 of the Final Rejection, the Examiner asserts that Marx discloses a barge-in switching unit. To the contrary, Marx, at column 7, lines 20-28 discloses software for detecting a barge-in. However, Marx does not suggest a barge-in switching unit for switching modes. Marx merely discloses that, when barge-in protection is provided, a prompt going out to the user should be stopped in response to sensing a barge-in. The section of Marx relied upon by the Examiner does not disclose a barge-in switching unit.
>
> Neither Polikaitis nor Marx (nor Bridges which was not applied for this purpose) disclose or fairly suggest switching modes, much less that the mode switching be achieved using a barge-in switching unit.

In response to the Appellant's argument in paragraph one and three above that requiring the barge-in switching unit now requires claim 11 to be interpreted in accordance with the Examiner's option (1) set forth in paragraph 17b of the Final Rejection, the Examiner cannot concur. It should pointed out that the Appellant has made an error in the above statement, where claim 12 is dependent upon claim 9 and therefore was interpreted that the Appellant is relying on claim 9 for option 1 to be interpreted. Whichever the case may be, the claim language of claim 12, uses terminology of "further comprising." Furthermore, there is no limitation in claim 12, which requires option 1 to be interpreted as there is no terminology that refers back to claim 9 as asserted by the Appellant. Hence, interpreting the barge-in switching unit independently as another component in the system is a reasonable interpretation.

In response to the Appellant's argument in paragraphs two and three above that Marx does not disclose a barge-in switching unit since discloses software for detecting a barge-in. To begin with, it should be noted that a barge-in switching unit for switching modes is not being claimed. Rather a barge-in switching unit is being claimed. The

claim does not provide any limitations to further define that functionality for the barge-in switching. Thus, Marx's barge-in software which is switching based on user speaking during a prompt reads on the claim 12. For example, the barge-in mode is active once the user speaks during a prompt and de-active when the user is not speaking, which is a switching mechanism (see Marx col. 7, lines 20-25). Further, the Appellant's published Specification provides description as to the barge-in switching unit being software in paragraph [0041], where the barge-in switching unit is a software switch.

Conclusion

For the above reasons, it is believed that the rejections should be sustained.

Respectfully submitted,

Paras Shah

03/24/2010

/David R Hudspeth/
Supervisory Patent Examiner, Art Unit 2626

Conferees:

David Hudspeth
/D.R.H./

James Wozniak

/James S. Wozniak/

Primary Examiner, Art Unit 2626

/Paras Shah/
Examiner, Art Unit 2626